



Lexical patterns in customary sayings

Mohammad Fadzeli Jaafar

*Programme of Linguistics, Faculty of Social Sciences and Humanities, Universiti Kebangsaan Malaysia, 43600 UKM
Bangi, Selangor, Malaysia*

ABSTRACT

Customary sayings are the traditional words delivered in the form of poetry and practiced in various social activities in Negeri Sembilan in marriage, family, criminal law and administration. Sudeen (1995) recorded that studies on customary sayings have been conducted from the early 20th century in the history of custom, marriage system, political system and the dispensation of inheritance. This study examines the content words and function words in customary sayings that have been categorized based on the semantic domain concept by Gliozzo (2006). Based on the content words and function words, this study has identified the customary words in the corpus data. The results show that more function words were used in the customary sayings. However, content words dominated the list of customary sayings, such as *Datuk* and *custom*. The function words '*nan*' and '*dek*' were found in all domains of the study namely, political, social and economic domains. These findings indicate that function words are not only grammatically functional, but also used to give an aesthetic impact through the articulation style of the local dialect. On the other hand, the analysis of the content words shows that the use of customary words is influenced by the environment and culture of Adat Perpatih.

Keywords: customary sayings; corpus linguistics; semantic domain; customary word; Adat Perpatih

First Received:

1 April 2017

Revised:

9 January 2018

Accepted:

9 March 2018

Final Proof Received:

26 May 2018

Published:

31 May 2018

How to cite (in APA style):

Nair, R. & Arshad, R. (2018). The discursive construction of teachers and implications for continuing professional development. *Indonesian Journal of Applied Linguistics*, 8(1), pp. 226-234. doi: 10.17509/ijal.v8i1.11461

INTRODUCTION

This study was motivated by customary sayings as the traditional words dealing with customary law (Ahmad, 1965), in which questions arise on the meaning of traditional words and customary laws. In terms of meaning, the word 'traditional' means "custom or belief that has existed for a long time" (Collins Cobuild English Dictionary, 1995:1776). This definition asserts that the term "traditional" also refers to a custom or a belief that has existed for a long time. Thus, traditional words in this context can be referred to as customary words. Specifically, customary words in the customary sayings are used to record the laws relating to custom, namely Adat Perpatih. Generally, the Malays in Peninsular Malaysia consider that there are two types of

Malay customs, namely Adat Perpatih and Adat Temenggung. The Adat Perpatih is practiced in Negeri Sembilan and Adat Temenggung is practiced outside of Negeri Sembilan namely, Melaka, Johor and Selangor (Ibrahim, 1993).

This study examines the use of customary words in sayings using the corpus linguistic approach. Based on this approach, a number of customary sayings data were generated in the corpus. The purpose is to obtain the frequencies or usage statistics of the content words and function words in the sayings. The content words and function words were clustered based on the domain. The study identifies the customary words, which are described according to the Adat Perpatih's context and system.

* Corresponding author:

Email: fadzeli@ukm.edu.my

The field of customary sayings has gained the attention of both Western and local researchers. Western researchers have documented the history of custom, marriage system, political system and the dispensation of inheritance since the beginning of the 20th century (Sudeen, 1995). Later, Sudeen recorded that the writings and documentations about customary sayings by local researchers had emerged in the 1950's. However, most of the previous studies focus on the documentation. The scientific study of customary sayings only emerged in the 90s, especially by Ibrahim (1993, 2003, 2010). In other words, the scientific research about customary sayings is still small in number (see Kadir, 1996; Sudeen, 1995). In this case, Kadir (1996) concluded the weaknesses of previous studies about sayings, i.e. (i) the documentations were not based on examination, description and systematic interpretation, (ii) the Western research lacks understanding of the social system and philosophies in the Adat Perpatih communities, (iii) the local studies are mostly ethnocentric. Therefore, it can be concluded that comprehensive and interpretative studies about customary sayings have not been conducted extensively.

With regard to the criticism and assessment of the previous studies, this study drew upon a number of customary sayings data collected by local researchers. In particular, this study examines the lexicon and their meanings in the customary sayings. The discussion will be based on the context or domain setting. The use of a large data and the study of the meaning of lexical in the sayings are expected to meet the nature of a comprehensive and an interpretative study.

METHOD

This study examines how the words were expressed in the customary sayings. To achieve this goal, as many as 6 customary sayings' collections were generated in the corpus data. The two aspects which were given attention in the method of this study are (i) the construction of corpus, and (ii) the construction of domain. The construction of corpus was done based on the linguistic corpus, while the construction of domain was based on the concept of semantic domain. The following are details on the methods of building the corpus and domain in this study.

This study has selected a number of customary sayings compilations as the corpus, making lexical as the study focus. The corpus data in this study will be used to answer the following questions: (i) what are the most widely used content words and function words in the sayings, (ii) what the customary words sayings are, and (iii) what the usage frequencies are. In other words, this study will highlight the lexical statistics studied, which is the basic technique in a corpus-based study (Biber, Conrad, & Reppen, 1998; Hunston, 2006). This finding will discover the lexical patterns in the sayings, thus, become the sayings' characteristic.

To date, there have been many lexical studies based on the corpus. For example, Jablonkai (2009)

studied lexical in newspapers and legislation texts, Esimaje (2012) examined the usage of lexical in sermons, Mazdayasna and Firouzi (2013) examined the lexical adjective in novels and technical texts; and Molavi, Koosha and Hosseini (2014) explored the lexicon in text books. These studies examined the lexical in various types of data. On the other hand, this study examines the customary sayings, which can be categorized as the literary genres

The procedure for the construction of corpus was conducted as follows: (i) collection of all data using the scanning equipment (scanner) through an Android application, namely the OCR Text Scanner, (ii) transferred into Word format, which is size 1 Word file, 665.2 kb (41,764 TOKENS words) which contained 417 pages, (iii) transfer of Word file to Plain Text format size 294.647 kb to generate the corpus using AntConc 3.4.4 w software (Windows) 2014 to obtain the frequency of content words and function words. For the purpose of this study, only the 10 highest content words and function words will be discussed.

Currently, the semantic domain is the latest development in the computational linguistic study. Related to this, the data of this study were clustered based on the semantic domain concept proposed by Gliozzo (2006, see also Gliozzo & Strapparava, 2009):

In our usage, Semantic Domains are common areas of human discussions, such as ECONOMICS, POLITICS, LAW, SCIENCE, which demonstrate lexical coherence. The Semantic Domain associated to a particular field is the set of domains specific terms belonging to it, and it is characterized by a set of *domain words* whose main property is to co-occur in texts.

The above statement mentions that semantic domain is often discussed in the domains of economics, politics, law and science, based on lexical relation. The semantic domain concept corresponds to the customary sayings as stated by Ibrahim (2010:1), "Adat Perpatih was built around a communal kin-based social, politic and economic system". This statement confirms that Adat Perpatih practiced among the Negeri Sembilan community was established based on the social, political and economic systems. Customary sayings have been descended from generation to generation and cover various aspects of life such as the appointment of leaders, matters of inheritance, law of marriage, social practice and the punitive system. All of these systematic practices can be divided into three main domains, namely, political, social and economic domains.

In actual fact, the domain concept was used by the local researchers, i.e. Selat (1975) and Ibrahim (2010). However, the usage of these domains was not discussed from the semantic aspect. Instead, they were using the 'system' concept; for instance, Selat (1975) states that each community has their own respective system. The Adat Perpatih has its own system manifested through customary sayings. In his research, Selat has categorized 6 sub-systems that exist in Adat Perpatih,

namely, value system, kinship system, political system, economic system, customary law system and stratification system. Ibrahim (2010) also discusses the Adat Perpatih system in Negeri Sembilan based on the following division namely, social, political and economic systems. The social system refers to descend, clan, adoption and marriage. The political system refers to the political hierarchy. The economic system is related to customary inheritance. Ibrahim's division is found to be better than Selat's categorization. At least, Ibrahim has established 3 main systems, namely social, political and economic systems. Each system is divided into several sub-systems, respectively.

However, based on the examination, it was found that the division of system and subsystem of the customary sayings by the local researchers shows an overlapping. For instance, the value system category by Selat is also found in other subsystems, for example political, economy, customary law and stratification systems. Ibrahim did not specifically discuss the customary sayings, instead, the Adat Perpatih system. His research focuses on the vanishing culture in Adat Perpatih based on the social, political and economic system. Thus, the categorization of customary sayings according to domain was not systematically conducted. Other weaknesses are both researchers, Selat and Ibrahim only used a few examples of customary saying in their discussions. Such studies are not comprehensive in nature, and the interpretation is somewhat limited to a specific example.

This study uses a new perspective in the language usage, i.e. considering the relevant association patterns. In this context, association patterns refer to the quantitative data, which will be linked to the interpretation (qualitative), which is the basic step in a corpus-based study (Biber, Conrad & Reppen, 1998). In other words, this study not only reports the frequencies of lexicon, but also explains the meaning of such lexicon in the corpus.

RESULTS AND DISCUSSION

As described above, the domain in this study pertains to the use of lexical criteria as the focus, i.e. the content words and function words. The content words refer to the main words in the grammar, namely nouns, verbs, adjectives and adverbs. While the function words refer to those which carry only grammatical meanings, namely conjunctions, relative pronouns and prepositions.

This study has identified the customary words in the content and function words list. The category of the customary words was done based on the book written by Ibrahim (2003) entitled *Glossary of Negeri Sembilan Custom*. He has documented almost 2600 Adat Perpatih's terms, sorted in alphabetical order. The terms collected consist of words, phrases, expressions, compounds and plural noun. For example:

- (1) Alam beraja
Luak berpenghulu
Suku berlembaga
Anak buah berbuapak
Orang semenda bertempat semenda

The state has a ruler
The district has a territorial chief
The clan has a clan chief
The sub-clan has an elder
The affined has affinal relative(s)

The above example (1) places the areas of governance of the head of custom, through the word *luak* (state), *penghulu* (territorial chief), *suku* (clan) and *buapak*. All these words are associated with the political system or hierarchy which exists in the Adat Perpatih starting from highest to lowest.

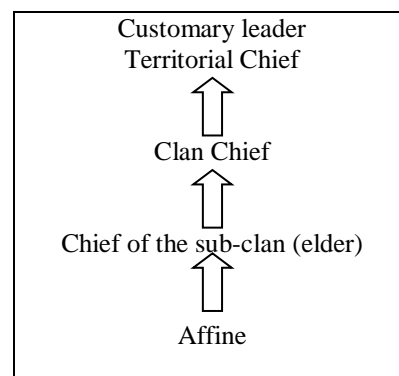


Figure 1: The structure of political leaders in Adat Perpatih.

In Adat Perpatih, the territorial chief is the supreme leader, and he is the head of a *luak* (state). The territorial chief is appointed by the chief of the clan that is the *lembaga* (clan head), followed by *buapak*, which has its own affine. This lexicon shares the same semantic domain, namely the same political system. Therefore, this lexicon is categorized as the political domain.

The social domain refers to the social events in the Adat Perpatih community. A total of customary sayings data about marriage customs, family and advice are categorized in the social domain. Based on observations, most of the customary sayings' data evolved on marriage rules or marriage law in Adat Perpatih. This will be discussed in the next section. The economic domain refers to any activities related to the cause of obtaining sustenance for the survival of Adat Perpatih community. This includes the management of environmental resources, such as land, crops and livestock.

A Comparison of the Content and Function Words.

Biber (2011) states that a corpus-based study is intended to obtain lexical distributions based on the analysis of 'keywords'. There are three keywords in this study, namely, content words, function words and customary

words. The discussion in this section focuses on the distribution of content words and function words in the corpus based on the domain of the study. Further, it will identify and describe the meaning of customary words in each domain stated.

As a whole, this corpus study consists of 41.764 words. The number is big compared to the previous researchers who studied only a few verses of the

customary sayings as examples of their studies. A comparison of the total number of domain words showed that the social domain is most dominant with 34,162 words, followed by the political domain with 7482 words, and the economic domains of only 120 words. Table 1 shows the distribution's presentation and description of the content words and function words for the whole data of the study.

Table 1. The overall frequencies of content words and function words

Content word	Frequency	Percentage (%)	Function word	Frequency	Percentage (%)
<i>datuk</i>	730	19.6	<i>nan</i>	792	17.2
<i>adat</i>	592	15.9	<i>yang</i>	630	13.7
<i>sembah</i>	435	11.7	<i>kok</i>	628	13.7
<i>orang</i>	430	11.6	<i>di</i>	546	11.9
<i>kata</i>	331	8.9	<i>ke</i>	469	10.2
<i>anak buah</i>	328	8.8	<i>sama</i>	355	7.7
<i>semenda</i>	258	6.9	<i>tak</i>	346	7.5
<i>tempat</i>	233	6.3	<i>pada</i>	288	6.3
<i>janji</i>	214	5.8	<i>dek</i>	274	6.0
<i>Mati</i>	167	4.5	<i>dan</i>	269	5.9
TOTAL	3718	100%	TOTAL	4597	100%

Table 1 displays the top 10 words for content words and function words. The word *datuk* and *adat* (custom) registered the highest frequency at 730 and 592. The next two words, *sembah* (obeisance) and *orang* (person) had a total of over 400 times. The word *kata* (word) and *anak buah* (affine) registered the usage of exceeding 300 times. The word *semenda*, *tempat* (place) and *janji* (promise) exceeded 200 times. The last one is the word *mati* (die) as much as 167 frequencies. From the list, 7 words are included in the noun, while 3 are verb (*sembah*, *janji*, *mati*).

The function words show that the relative pronouns of *nan* and *yang* (as) registered the highest total number of 1422 times (over 30%). The word *kok* (if) is the third highest. The prepositions *di* (at) and *ke* (to) have high frequencies of occurrence exceeding 400 times. They are followed by *sama* (same) and the negative conjunctive *tak* (no), exceeding 300 times. The last three words are the preposition *pada* (at), *dek* (because) and *dan* (and). Based on this list, preposition registered the highest usage, a total of five words, *di*, *ke*,

sama, *pada*, *dek* (at, to, same. because). Followed by 2 relative pronouns *nan*, *yang* (which) while the adverb word *kok* (if), negative conjunctive *tak* (no) and the conjunction *dan* (and) once of each respectively.

Based on the list of customary words in the *Negeri Sembilan Custom's Glossary* by Ibrahim (2003), there are 8 words in the content word category that is, the customary words which registered a percentage of 83.8%. While only 3 function words namely, *nan*, *kok* and *dek* (which, if, because) are in the customary word category (36.9%) respectively. These findings indicate that the usage of customary words is more common in the content word category.

Social Domain

On the whole, the social domain registered the most number of lexical than the political and economic domains (see Table 2). This gives the impression that customary sayings are highlighting more of the social aspects.

Table 2: Frequency of content words and function words in the social domain

Content Word	Frequency	Percentage (%)	Function Word	Frequency	Percentage (%)
<i>datuk</i>	678	21.5	<i>nan</i>	726	19.2
<i>adat</i>	409	13.0	<i>kok</i>	550	14.6
<i>sembah</i>	399	12.7	<i>yang</i>	509	13.5
<i>orang</i>	363	11.5	<i>di</i>	407	10.8
<i>kata</i>	273	8.7	<i>ke</i>	351	9.2
<i>anak buah</i>	271	8.6	<i>sama</i>	303	8.0
<i>semenda</i>	230	7.3	<i>tak</i>	273	7.2
<i>janji</i>	195	6.2	<i>dan</i>	233	6.2
<i>tempat</i>	181	5.7	<i>pada</i>	214	5.7
<i>hari</i>	151	4.8	<i>dah</i>	213	5.6
TOTAL	3150	100%	TOTAL	3779	100%

It is noted that the usage of content words and function words for social domain is not much different from the overall data in table 1. Only function words show a slight difference due to the absent of the word

dek (because). The word *datuk* and *adat* (custom) still registered the highest number of frequency. Similarly, to the word *nan* and *yang* (as) for the function words.

The discussion in this section focuses on the lexical meaning used in the social domain based on the corpus. For example, the word *Datuk* which was used the most has a few references. In the *Negeri Sembilan Custom Glossaries*, Ibrahim, (2003) divides at least three meanings of *datuk*, namely: i) the noble person, (ii) the father of a person's father, (iii) title of the Customary Head with the level of Undang, Territorial chief, Clan Head and Buapak. Based on the meanings stated here, the third meaning of *datuk* is most suitable for the Adat Perpatih system. For example,

- (2) Menjemput anak buah Datuk
Inviting the Datuk's affine

The example above uses a special name for the word *Datuk*. Almost all of the entire data wrote the word *Datuk* with a proper noun. This shows the impression that the word *Datuk* refers to a title or a particular name related to the Adat Perpatih system. Part of the data also used the phrase *anak buah Datuk* (Datuk's affine). In this context, *anak buah* (affine) refers to a member or members of the family cluster headed by the *Datuk*. In the Adat Perpatih system, *Datuk* refers to the territorial chief or it also refers to custom *Datuk* or the custom head in an area.

On the other hand, customary lexical refers directly to the rules, law or the law of the Adat Perpatih system. It is anticipated that the usage of *adat* (custom) word is one of the most used, particularly in the social and political domains because both domains are more likely implying the regulation. Typical words in sayings which uses the *adat* (custom) word are:

- (3) Adat bersendikan hukum
Custom is articulated by the Law

Asserted in the example above, the customs practiced among the customary community is in accordance with the Islamic law or religion. This is true in the social domain, for example, the customary marriage. Marriage is part of the rules of life. The Adat Perpatih community appreciates living together (Selat, 1975). Thus, customary sayings were made as a tool for describing marriage customs.

The data show that the word *orang* (person) as the fourth highest in the social domain. Basically, the word *orang* (person) is general in nature, namely referring to the community. The data shows that *orang tua* (elderly affine) and *orang muda* (young affine) refer to the different social groups. The word *Orang semenda* (semenda person) was also found in the data. In Adat Perpatih, the word *Orang semenda* (semenda person) refers to the affine who come into the clan other than his clan through marriage (Manaf, 2015). In the Adat Perpatih system, the men are required to live in the wife's area (matrilineal).

The *anak buah* (affine) lexical is often associated with one of the clan's name. The phrase *anak buah* (affine) refers to a member (i) a territorial chief/undang; the affine or the population of the entire luak (ii) to the Clan head: clan's members lead regardless of their places of residence. (iii) to Buapak: Members of the *perut* are led regardless of their places of residence (iv) Besar: Members of an area are led regardless of their places of residence (v) to Kadim: members of a place are led regardless of their places of residence.

Political Domain

The following discussion and Table 3 show the statistics for political domain based on the overall data of the study.

Table 3: Frequency of content words and function words in the political domain

Content Word	Frequency	Percentage (%)	Function Word	Frequency	Percentage (%)
<i>adat</i>	183	30.0	<i>di</i>	132	15.0
<i>orang</i>	67	11.1	<i>ke</i>	118	13.3
<i>kata</i>	58	9.5	<i>yang</i>	118	13.3
<i>anak buah</i>	57	9.3	<i>nang</i>	113	12.8
<i>Datuk</i>	52	8.5	<i>kok</i>	78	8.8
<i>tempat</i>	52	8.5	<i>pada</i>	74	8.4
<i>raja</i>	39	6.4	<i>tak</i>	70	7.9
<i>sembah</i>	36	5.9	<i>dek</i>	65	7.4
<i>mati</i>	35	5.7	<i>nan</i>	64	7.2
<i>tanah</i>	31	5.1	<i>sama</i>	52	5.9
TOTAL	610	100%	TOTAL	884	100%

Similar to the social domain, the content word in the political domain likely to be used is the word *adat* (custom) with the highest frequency of 183 times (30%). The word *orang* (person) is the second highest registered at 67 times (11.1%). The word *kata* (word) and *anak buah* (affine), each of which, amounts to almost the same total number. They are followed by the words *datuk* and *tempat* (place), which are the same total number of 52 times (8.5%). The word *raja* (King) occurred 39 times (6.4%), which is not much different

from the last 3 words, namely *sembah*, *mati* dan *tanah* (obeisance, death, land).

In the leadership context, customary sayings reveal that every member of the society has their respective roles. For example,

- (4) Adat berkampung, kampung pertahankan
Adat bernegeri, negeri pertahankan
Adat berpemimpin, pemimpin pertahankan

The norm of a village, the village must be defended

The norm of a state, the state must be defended

The norm of having a leader, the leaders need to defend

Example (4) above explains that custom functions in various contexts, i.e. as the norm of a village, the village should be defended. As the norm of a state, then the state must be defended, if there is a leader then the leader should defend. If the custom is in the leadership context, then the appointed leader needs to defend. The *adat* (custom) lexical in the political domain is often associated with the custom's chief or leader according to the order. For example,

- (5) Raja berkeadilan,
Penghulu / undang ibu adat,
Lembaga yang punya adat,
Buapak menjalankan adat,
Orang semenda menjaga dan memelihara adat,
Anak buah memakai adat.
- A just King,
The Territorial chief/Undang is the 'mother of the custom',
The Clan head owns the custom,
Buapak implement the custom,
The semenda affine keep and maintain the custom,
The affine practice the custom.*

The above sayings show that the King functions as a judge or referred to as 'justice'. In the Adat Perpatih, the King has no power over any customary *luak* in Negeri Sembilan except within his area namely, Seri Menanti which is specified for him. Seri Menanti is the Royal Town of Negeri Sembilan, which is also the place of residence for the Yang Di-Pertuan Besar of Negeri Sembilan. The King or Yang di-Pertuan Besar is only a symbol of justice and functions to preserve unity in Negeri Sembilan. On the other hand, the territorial chief is the highest head or leader of the custom. Therefore, it is referred to as the 'mother of the custom'. The clan head is the leader of a kinship group called clan. It is followed by a *buapak* as the head for the group called the *perut* (sub-clan). Buapak is appointed through consensus by the affines and agreed by the *Lembaga* (clan head). Therefore, buapak serves as a custom implementer. The *semenda* affine is assigned to take care of the custom, and the affines are as the custom's practitioners.

Similar to the social domain, function words for the political domain has almost the same pattern. The different is that the word *di* (at) registered the highest frequency of 132 times (15.0%) respectively, followed by two function words, *ke* (to) and *yang* (as), with the same frequency of 118 times (13.3%). The word *nang* occurred 113 times (12.8%) compared to the word *nan* which occurred 64 times (7.2%) respectively. Both

these words are different in form but have the same function. Other function words, such as *kok* (if), *pada* (on) and *tak* (no), amount to 70-78 times. Finally, the word *dek* (because) and *sama* (same) occurred 65 times (7.4%) and 52 times (5.9%).

Based on the corpus, the conjunction *di* (at) and *ke* (to) serve to indicate a place, for example:

- (6) Raja bertitah di istananya
Pengkulu berundang di balainya
- The King utters in his castle
The territorial chief practice in his court*
- (7) Ke bukit sama didaki
Ke lurah sama dituruni
- To the hill together we climb
To the valley together we descend*

Example (6) describes that the King will utter his command in the palace, while the territorial chief will implement the law in his court. Example (7) shows the concept of consensus in the society, namely enduring the hardship together. Similar to *di* (at), the word *ke* (to) functions to show a place, to the hill (a high place) and to the valley (a low place).

The interesting conjunction words to be discussed here are the relative pronoun variations *nang*, *yang* and *nan* (as). Following are examples of the usage of all the relative pronouns in the data:

- (8) Susun jari **nang** sepuluh
Angkat sembah pada Datuk
Kampung nang bersudut
Sawah nang berlopak
- The ten fingers are arranged
Obeisance the Datuk
Village with corners
Paddy fields with puddles*
- (9) Kok **nan** banyak ditambah
Kok **nan** lebih dicari
- If it is a lot it will be added
If it is more keep seeking*
- (10) Yang buruk diperbaharui
Yang usang diganti
- The old one is refurbished
The dilapidated is replaced*

The examples above show that the words *nang*, *nan* and *yang* (as) differ only in the form but have the same function which can be mutually replaced in sentences. For example, *nang* in example (8) can be replaced with *nan* and *yang*. Based on research, this study found that the word *nang* is an influence from the

Minangkabau language. However, this is not surprising because the Adat Perpatih system originates from Minangkabau (Ibrahim 1993). Subsequently, it was discovered that although the forms are alternate, the corpus shows that only the elements of *yang* and *nan* can be present at the beginning of a sentence. This reveals that the usage of the word *nan* is restricted as a relative pronoun in a sentence.

Further discussion on the function word *kok* (if) and *dek* (because), which is included in the category of customary words. For example:

- (11) **Kok** tua dimuliakan
Kok muda dikasihi

If it is the elderly is respected
If it is the youngsters are loved

- (12) Adat yang tak lapuk **dek** hujan
Tak lekang **dek** panas

An ever-enduring custom

Both the words *kok* (if) and *dek* (because) belong to the Negeri Sembilan dialect style of pronunciation. In the Malay language, the standard word *kok* (if) means *kalau*, while the word *dek* (because) refers to *oleh* (by) or *kerana* (because). Based on example (11) the word

kok (if) should be replaced with the conjunction word *kalau* (if), while the word *dek* (because) can be replaced with the conjunction *kerana* (as). The meaning of sayings in example (11) is that both the young and old must respect each other, while example (12) insists that the custom practiced is suitable at all times. One reason is that the Adat Perpatih as aforementioned is based on the Islamic law. However, the data of this study show that the word *kok* (if) is only present at the beginning of a sentence, while the word *dek* (because) is in the middle of the sentence. In other words, the word *kok* (if) is used at the beginning of a sentence, and the word *dek* (because) is used as a conjunction between phrases in a sentence.

Economic Domain

As described, the economic domain is the smallest total amount compared to the two domains above. Based on the examination, the customary sayings are highlighting more of the rules based on the Adat Perpatih community's practices for example, the marriage rules in social domain. On the other hand, the economic domain is more focused on the Adat Perpatih's economic activities for example, agriculture. Table 4 displays the presentation of the frequency data of content words and function words in the economic domain.

Table 4. Frequency of content words and function words in the economic domain

Content Word	Frequency	Percentage (%)	Function Word	Frequency	Percentage (%)
<i>elok</i>	4	13.80	<i>dek</i>	7	22.58
<i>hilang</i>	4	13.80	<i>di</i>	7	22.58
<i>air</i>	3	10.34	<i>segala</i>	4	12.90
<i>emas</i>	3	10.34	<i>tak</i>	3	9.68
<i>majlis</i>	3	10.34	<i>yang</i>	3	9.68
<i>nilai</i>	3	10.34	<i>nan</i>	2	6.45
<i>padi</i>	3	10.34	<i>tepi</i>	2	6.45
<i>bahasa</i>	2	6.90	<i>kerana</i>	1	3.23
<i>bangsa</i>	2	6.90	<i>lagi</i>	1	3.23
<i>berarak</i>	2	6.90	<i>tentang</i>	1	3.23
Total	29	100%	Total	31	100%

The use of content words in the economic domain shows a significant difference compared to the political and social domains. For example, based on table 3, none of the lexical is the same as the two previous domains. A comparison of the total number of words was found to be insignificant because the frequency is only between 2 and 4. Likewise, the function words occurred with the frequency of 1 to 7. From the words' class aspect, a large number of the content words are nouns, except the adjective word *elok* (nice), and the verb *hilang* (lost) and *berarak* (scud). Although their numbers are small, it was detected that some words are associated directly with the customary words, for example *air* (water), *emas* (gold), *majlis* (council) and *bangsa* (race). These words are used as metaphors taken from the environment surrounding the Adat Perpatih community. Although the word *padi* (paddy) is not included in the customary words, yet it is significant as

an economic source of the Adat Perpatih community and is an essential component in the customary inheritance (Ibrahim, 2010). It is similar to the word *air* (water) which helps in the process of paddy planting.

Comparison of Customary Words According to Domain

In this section, the use of customary words according to domain is discussed. The comparison will show the dominant form of customary words involving the content and function words (see Table 5 for detailed comparison).

Based on the comparison, the social domain has the highest number of customary words, at 3987, followed by 528 words for the political domain and only 26 words in the economic domain. This is expected because the social domain has the most number of words.

Table 5. Frequency of customary words according to domain

Social Domain			Political Domain			Economic Domain		
Customary Word	Frequency	Percentage	Customary Word	Frequency	Percentage	Customary Word	Frequency	Percentage
<i>nan</i>	726	(18.63%)	<i>adat</i>	183	(34.65%)	<i>dek</i>	7	(26.92%)
<i>datuk</i>	678	(17.40%)	<i>nang</i>	57	(10.79%)	<i>air</i>	3	(11.54%)
<i>kok</i>	550	(14.11%)	<i>anak buah</i>	57	(10.79%)	<i>emas</i>	3	(11.54%)
<i>adat</i>	409	(10.50%)	<i>kok</i>	52	(9.84%)	<i>majlis</i>	3	(11.54%)
<i>orang</i>	363	(9.31%)	<i>orang</i>	39	(7.38%)	<i>beremas</i>	2	(7.69%)
<i>kata</i>	273	(7.01%)	<i>dek</i>	36	(6.81%)	<i>hari</i>	2	(7.69%)
<i>anak buah</i>	271	(6.95%)	<i>nan</i>	31	(5.87%)	<i>lenggang</i>	2	(7.69%)
<i>semenda</i>	230	(5.90%)	<i>kata</i>	28	(5.30%)	<i>nan</i>	2	(7.69%)
<i>dek</i>	202	(5.18%)	<i>datuk</i>	24	(4.54%)	<i>akar</i>	1	(3.85%)
<i>janji</i>	195	(5.00%)	<i>tempat</i>	21	(3.97%)	<i>kain</i>	1	(3.85%)
Total	3,897			528			26	

The social domain shows that the function word *nan* has the highest percentage of usage at 726 times (18.63%) respectively. In addition, there are also other function words in the customary word category which is the ten highest, namely *kok* (if), a total of 550 times (14.11%), and *dek* (because), a total of 202 times (5.18%) respectively. This discovery is surprising because the function words apparently contributed to the formation of the customary sayings. In fact, the function words also have a high number in the political domain, for example *nang* (as), as many as 57 times (10.79%), *kok* (if), 52 times (9.84%), *dek* (because), 36 times (6.81%) and *nan* (as), 31 times (5.87%). In the economic domain, the two function words are in the 10 highest are *dek* (because) and *nan* (as).

If the three domains are examined, the function words were significant in the customary sayings when used in all domains, for example, *nan* (as) and *dek* (because). The word *kok* (if) is used in the social and political domains. While the word *nang* (as) is only used in the political domain. This has shown interesting lexical usage patterns in the customary sayings studied.

For the content words, there are a number of customary words which are often used in the social and political domains, namely *datuk*, *adat* (custom), *orang* (person), *kata* (word) and *anak buah* (affine). The rest of the words such as *semenda* and *janji* (promise) are only used in the social domain, while the word *tempat* (place) was found only in the political domain. In the economic domain, the noun dominated the usage of customary words, namely *air* (water), *emas* (gold), *majlis* (council), *hari* (day), *akar* (root) and *kain* (cloth). Only two verbs are used in this domain namely, *beremas* (golden) and *lenggang* (swing).

CONCLUSION

The corpus linguistics approach has successfully managed a large data and thus, systematically describing the customary lexical patterns. With this, the use of computing technology has been beneficial in the study of literature texts. The semantic domain concept has contributed to lexical clustering according to the respective domains. The semantic domain information

has facilitated the process of identifying the content words, function words, and customary words.

Each customary word used in the sayings has a value or meaning related to the customary community. Based on all the customary words discussed, customary words are divided into two types, namely, (i) specific customary words and (ii) general customary words. A specific customary word refers to the word that can be directly related to the customary system, such as *Datuk*, *anak buah* (affine), *adat* (custom), *semenda*, *raja* (King), *majlis* (council), *kok* (if), *dek* (because), *nan* and *nang*(as). Special customary words are found to be accentuating more of the traditional and archaic forms. The general customary words are words which are not directly related to customs, such as *orang* (person), *tempat* (place), *emas* (gold), *hari* (day), *kain* (cloth) and *akar* (root). In terms of statistics, special customary words are more dominantly used compared to the general customary words.

Based on the word class analysis, it was found that the customary words consist both the content and function words. More interestingly is that function words dominate the social and political domains. Analysis of the function words in customary sayings has never been discussed by the previous researchers. The data indicates that the function words are not just grammatically functioning in a sentence. For example, the word *kok if*, will not only serve as the adverb, but deliberately repeated at the beginning of a line to create an aesthetics effect through parallelism elements, with the local style pronunciation. On the other hand, the word *kok* (if), *dek* (because), and *nan/nang* (as) show the existence of the local dialect influence (Negeri Sembilan dialect) and Minangkabau dialects. Therefore, it is not surprising that function words are also the traditional words or part of the customary words.

Analysis of the meaning shows that customary words can be categorized into three, namely, peoples' name (*Datuk*, *orang* person, *anak buah* affine and *semenda*), abstract nouns (*adat* custom, *majlis* council and *hari* day) and object nouns (*akar* root, *kain* cloth, *emas* gold and *air* water). This case provides an overview of the customary words making use of all language elements existing in the environment and culture which surrounds Adat Perpatih.

ACKNOWLEDGEMENT

This paper is generated from research project coded FRGS/1/2015/WAB04/UKM/02/2 funded by Ministry of Higher Education, Malaysia.

REFERENCES

- Ahmad, Z. A. (1965). *Ilmu Mengarang Melayu*. Kuala Lumpur: Dewan Bahasa dan Pustaka.
- Biber, D. (2011). Corpus linguistics and the study of literature: Back to the future? *Scientific Study of Literature*, 1(1), 15-23.
- Biber, D., Conrad, S., Reppen, R. (1998). *Corpus linguistics: Investigating language structure and use*. Cambridge: Cambridge University Press.
- Collins Cobuild English Dictionary. (1995). London: HarperCollins Publisher.
- Esimaje, A. U. (2012). A Corpus-based Lexical Study of Sermons in Nigeria. *English Language Teaching* 5(9), 24-32.
- Gliozzo, A. (2006). Semantic domains and linguistic theory. In *Proceedings of the LREC 2006 workshop: Toward Computational Models of Literary Analysis*. Genova, Italy.
- Gliozzo, A., & Strapparava, C. (2009). Semantic domains. In *Semantic Domains in Computational Linguistics* (pp. 13-32). Springer, Berlin, Heidelberg.
- Hunston, S. (2006). 'Corpus Linguistics'. In K. Brown, *Encyclopedia of Language and Linguistics* (2nd ed.) (pp. 234-248). Amsterdam: Elsevier.
- Ibrahim, N. H. (1993). *Adat perpatih; perbezaan dan persamaannya dengan Adat Temenggung*. Kuala Lumpur: Penerbit Fajar Bakti Sdn. Bhd.
- Ibrahim, N. H. (2003). *Glosari adat Negeri Sembilan*. Selangor: AFY Communications.
- Ibrahim, N. H. (2010). *Vanishing culture of the adat perpatih*. Retrieved from <http://www.malaycivilization.com.my/artikel-padat/>
- Jablonkai, R. (2009). "In the light of": A corpus-based analysis of lexical bundles in two eu-related registers. *WoPaLP*, 3, 1-27.
- Kadir, K. A. (1996). *Perbilangan adat: penilaian hasil penyelidikan*. Kertas kerja Seminar Pengemaskinian Pensejarahan Negeri Sembilan.
- Manaf, A. A. (2015). Peranan dan fungsi orang Semenda dalam usaha pemeraksanaan Tanah Adat. In M. Yusof, M. A. Mis, Y. Yunos, & M. R. Razak (Eds.), *Malaysia – Indonesia: Bahasa, Sastera dan Budaya*. Bangi: Penerbit Universiti Kebangsaan Malaysia.
- Mazdayasna, G., & Firouzi, M. (2013). A corpus based study of adjectives in literary and technical texts. *EFL Journal*, 2(1), 73-88.
- Molavi, A., Koosha, M., & Hosseini, H. (2014). A comparative corpus-based analysis of lexical collocations used in EFL textbooks. *Latin American Journal of Content and Language Integrated Learning*, 7(1), 66-81.
- Selat, N. (1975). *Sistem sosial adat perpatih*. Kuala Lumpur: Utusan Publications & Distributors Sdn. Bhd.
- Sudeen, A. (1995). Penulisan perbilangan adat: satu sorotan awal. *Warisan*, 23, 15-29.

Corpus data

- Caldecott, A., & Winstedt, R. O. (1918). Jebebu Customary Songs and Sayings. *Journal of the Straits Branch of the Royal Asiatic Society*, (78), 3-41.
- Ibrahim, N. H. (1996). *Perkahwinan adat di Negeri Sembilan*. Shah Alam. Fajar Bakti.
- Saluddin, M. R. (2007). *Anjung teromba*. Pulau Pinang: Goal Intelligent Publishing Sdn. Bhd. cor
- Saluddin, M. R. (2007). *Hari-hari mendatang Adat Perpatih*. Pulau Pinang: Goal Intelligent Publishing Sdn. Bhd.
- Saluddin, M. R. (2009). *Teromba dalam puisi melayu tradisional*. Selangor: Karisma Publication Sdn Bhd.
- Saluddin M. R. (2009). *Teromba sebagai alat komunikasi dalam kepimpinan adat perpatih*. Selangor: Karisma Publication Sdn Bhd.